1 Homework 9

Show that for any $\rho \in \mathbb{C}^{d \times d}$, there exists quantum channels $C : \mathbb{C}^{d \times d} \to \mathbb{C}$ and $D : \mathbb{C} \to \mathbb{C}^{d \times d}$ such that:

$$\Delta(D(C(\rho)), \rho) = 0.$$ 

2 Shannon Channel Coding Theorem

The goal of this tutorial is to prove Shannon channel coding theorem. First, recall the definition of a code and the capacity of a channel:

**Definition 2.1.** A $(n, R, \delta)$ code for the channel $W = \{ W(y|x) \}_{x \in X, y \in Y}$ is a pair $E, D$ such that:

1. $E : \{0, 1\}^{Rn} \to X^n$,
2. $D : Y^n \to \{0, 1\}^{Rn}$,
3. With $x^n = x_1 \ldots x_n$, $y^n = y_1 \ldots y_n$ and $W^n(y^n|x^n) := W(y_1|x_1)W(y_2|x_2)\ldots W(y_n|x_n)$:

$$\frac{1}{2^{Rn}} \sum_{s \in \{0, 1\}^{Rn}} \sum_{y^n \in Y^n : D(y^n) = s} W^n(y^n|E(s)) \geq 1 - \delta.$$ 

It describes the average over all messages $s$ of the probability of successfully decoding $s$, using $n$ independent copies of the channel $W$.

**Definition 2.2.** For a given channel $W = \{ W(y|x) \}_{x \in X, y \in Y}$, define the capacity of $W$ by:

$$C(W) = \max_{P_X} I(X : Y),$$

where the joint distribution over $X, Y$ is defined by $P_{XY}(x, y) = P_X(x)W(y|x)$.

**Theorem 2.3** (Shannon Channel Coding Theorem). For $R < C(W)$, there exists a sequence of $(n, R, \delta_n)$ codes for $W$ with $\delta_n \to 0$.

2.1 The decoder

We will assume that $R < C(W)$ is fixed. Let $P_X$ achieving the maximum in the definition of $C(W)$, and define $P_{XY}(x, y) = P_X(x)W(y|x)$. We will first assume that the encoder $E$ is given:

1. What is the best choice for $D$?

However, this expression is hard to analyse. We will rather use the following decoder:

$$D(y^n) = \begin{cases} s & \text{if there is a unique } s \text{ such that } W^n(y^n|E(s)) \geq \alpha(n, y^n), \\ s_0 & \text{otherwise.} \end{cases}$$

where $\alpha(n, y^n)$ will be defined later.
2. Give an expression for $P_{\text{err},s}$, the probability of error for message $s$.

3. Prove that $P_{\text{err},s} \leq P_{\text{err},s}^1 + P_{\text{err},s}^2$, with:

$$P_{\text{err},s}^1 := \sum_{y^n \in \mathcal{Y}} W^n(y^n|E(s))1_{W^n(y^n|E(s)) < \alpha(n,y^n)}$$

$$P_{\text{err},s}^2 := \sum_{y^n \in \mathcal{Y}} W^n(y^n|E(s)) \sum_{s' \neq s} 1_{W^n(y^n|E(s')) \geq \alpha(n,y^n)}$$

### 2.2 The encoder

We will use the probabilistic method to choose the encoder. For any message $s$, we will take $E(s) = x_1 x_2 \ldots x_n$, where all $x_i$ are chosen independently following the law $P_X$. The global encoding scheme is $E$ where all $E(s)$ are chosen independently following the previous distribution.

Our objective is to show that $\mathbb{E}_E[P_{\text{err}}] \to 0$, where $P_{\text{err}} = \frac{1}{2^m} \sum_{s \in \{0,1\}^n} P_{\text{err},s}$.

1. How can you prove Theorem 2.3 if you have $\mathbb{E}_E[P_{\text{err}}] \to 0$?

2. Let us take now $\alpha(n,y^n) = K(n,\varepsilon) P_{Y^n}(y^n)$, where $K(n,\varepsilon)$ will be defined later, with:

$$P_{Y^n}(y^n) = \sum_{x^n \in \mathcal{X}^n} P_{X^nY^n}(x^n, y^n) = \sum_{x^n \in \mathcal{X}^n} P_{X^n}(x^n) W^n(y^n|x^n) .$$

(a) With iid. variables $X_i Y_i$ following distribution $P_{XY}$, show that:

$$\mathbb{E}_E[P_{\text{err},s}^1] = \mathbb{P}\left(\prod_{i=1}^{n} W(Y_i|X_i) < K(n,\varepsilon) \prod_{i=1}^{n} P_Y(Y_i) \right) .$$

(b) Define $i_{XY}(x,y) := \log\left(\frac{P_{XY}(x,y)}{P_X(x)P_Y(y)}\right)$. What is the value of $\mathbb{E}[i_{XY}(X_i, Y_i)]$?

(c) Show that:

$$\mathbb{E}_E[P_{\text{err},s}^1] = \mathbb{P}\left(\sum_{i=1}^{n} i_{XY}(X_i, Y_i) < \log(K(n,\varepsilon)) \right) .$$

(d) Using the weak law of large numbers\(^1\), give some sufficient conditions on $K(n,\varepsilon)$ to have $\mathbb{E}_E[P_{\text{err},s}^1] \to 0$.

3. Give an upper bound on $\mathbb{E}_E[P_{\text{err},s}^2]$ depending on $K(n,\varepsilon)$, and give some sufficient conditions on $K(n,\varepsilon)$ to have $\mathbb{E}_E[P_{\text{err},s}^2] \to 0$.


### 2.3 An application

1. Compute $C(W)$ for the bit flip channel $W$, ie. $W(b|b) = 1 - f$, $W(\overline{b}|b) = f$ for $b \in \{0,1\}$ and $f \in [0,1]$.

\(^1\text{If } X_i \text{ are iid., then } \mathbb{P}\left(\frac{1}{n} \sum_{i=1}^{n} X_i - \mathbb{E}[X_i] < \varepsilon\right) \to 1\)